

로보스트 방법을 이용한 EEG 신호의 전력밀도 추정

김택수^o, 허재만, 김종순, 유선국*, 박상희

연세대학교 공과대학 전기공학과

순천향대학교 공과대학 전기공학과

Power Spectrum Estimation of EEG Signal Using Robust Method

T.S.Kim^o, J.M.Huh, J.S.Kim, S.K.Yoo*, S.H.Park

Dept. of Electrical Eng., Yonsei University

Dept. of Electrical Eng., Soonchunhyang University

ABSTRACT

EEG(Electroencephalogram) background signals can be represented as the sum of a conventional AR(Autoregressive) process and an innovation process, or a prediction error process. We have seen that conventional estimation techniques, such as least square estimates(LSE) or Gaussian maximum likelihood estimates(MLE-G) are optimal when the innovation process satisfies the Gaussian or presumed distribution. But when the data are contaminated by outliers, or artifacts, these assumptions are not met and conventional estimation techniques can badly fail and be strongly biased. It is known that EEG can be easily affected by artifacts. So we suggest a robust estimation technique which considerably performs well against those artifacts.

1. 서론

시계열(time series) 데이터의 주파수 성분, 즉 전력밀도 스펙트럼(Power Density Spectrum)을 추정할 때, 주로 Maximum Likelihood Method나 최소자승법(Least Square Method)과 같은 고전적인 통계적 추정기법을 사용한다. 이들 방법은 데이터의 분포를 미리 알고 있거나, 혹은 데이터가 어떤 특정한 분포를 하고 있다고 가정을 한다. 따라서 데이터가 이러한 가정을 만족하는 분포를 갖는 경우에는 위의 두 가지 방법을 이용한 스펙트럼 추정이 적절하다.

그러나 실제로 여러가지 이유로 데이터가 가정된 분포를 벗어나는 경우가 자주 발생하게 되는데, 이 때는 Maximum Likelihood Method나 최소자승법과 같은 방법으로는 스펙트럼 추정을 제대로 할 수 없다. 가정된 분포를 벗어난 데이터를 통계적 용어로 아웃라이어(outlier)라고 하는데, 아웃라이어가 존재하면 스펙트럼에 왜곡이 생겨 특징을 보려고 하는 최고치(peak)가 상실된다. 이러한 아웃라이어의 영향을 줄이려면 데이터 수를 상당히 증가시켜야 하는데 이렇게 하면 그 만큼 계산시간이 증가되어 비효율적이다.

2. Outlier와 Robustness

통계적 모델링 방법에서는 추정오차가 가우시안 분포(Gaussian distribution)를 한다고 가정하는 경우가 많다. 그러나 실제의 데이터는 아웃라이어와 같은 비정상적인 데이터를 포함하고 있는 경우가 빈번하기 때문에 이러한 가정으로는 주어진 데이터를 항상 정확히 분석하기 어렵다. 그리고 이러한 아웃라이어는 가우시안 분포와 같이 미리 가정한 분포와 일치하지 않는다. 대개의 경우 아웃라이어는 그림 2.1과 같이 중앙 부분은 가우시안 분포와 비슷하지만 꼬리부분에서는 정상 가우시안 분포보다 두껍거나(long-tailed, 혹은 heavy-tailed), 얇은(truncated, 혹은 short-tailed)분포를 갖는다.[2]

본 연구에서 처리하고자 하는 배경뇌파 신호는 여러가지 원인에 의하여 아티팩트(artifact)가 생기는데, 이는 외부의 전자기파 간섭, 전극의 접촉저항의 변화, 눈의 깜박임이나 동공의 움직임, 근육이나 심장전위 등이다.[1] 이러한 아티팩트들은 가정된 분포를 벗어난 아웃라이어로서 작용하게 되어, 기존의 통계적 방법으로는 스펙트럼을 정확히 추정할 수 없다.

따라서 본 연구에서는 이러한 아티팩트에 의하여 오염된 배경뇌파 신호의 전력밀도 스펙트럼을 추정할 때 생기는 아웃라이어의 문제점을 해결하기 위하여 건실성(Robustness)의 개념을 도입하였다. 로보스트 방법은 가정된 분포를 만족하는 데이터에 대해서는 기존의 방법 즉, Maximum Likelihood Method나 최소자승법과 같은 결과를 얻을 수 있고, 아웃라이어에 의하여 오염된 데이터에 대해서도 왜곡현상이 없는 건실한 스펙트럼을 추정할 수 있는 방법이다.

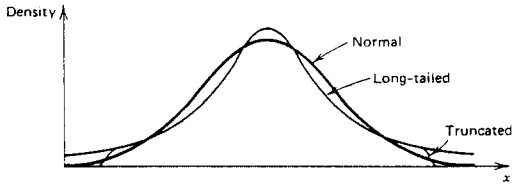


그림 2.1 아웃라이어의 분포곡선
Fig. 2.1 Distribution curve of outliers

아웃라이어는 그 스케일이 관측데이터의 스케일에 비하여는 크지 않더라도, 예측 오차(prediction error, 혹은 innovation process)에 비하여 어느 정도 커지면, 전력밀도 스펙트럼에 상당한 영향을 미치게 된다. 이러한 아웃라이어들에 대처하기 위하여 로버스트 스펙트럼 추정방법과 같은 특별한 방법의 필요성이 대두된다.

로버스트 추정(robust estimator)을 통계적으로 정의하면 다음과 같다.[3]

'F'를 관측 데이터의 정상분포 또는 가정된 분포라고 하고, $\{G_i\}$ 를 F와 '비슷한' 분포를 갖는 관측 데이터의 실제 분포들의 집합이라고 하자. 이 때 $\{F\} \cup \{G_i\}$ 집합 전체에 해당하는 분포에 대하여 스펙트럼의 추정을 재대로 할 때, 추정 T를 로버스트 추정이라고 한다. 일반적으로 F는 가우시안 분포에 해당하고, $\{G_i\}$ 는 두꺼운 꼬리를 갖는 (heavy-tailed)분포에 해당한다.

3. Robust Filtering Algorithm

로버스트 필터링 알고리즘은, 우선 로버스트 방법으로 데이터를 프리화이트닝(whitening)한 다음, 고전적 방법으로 스펙트럼을 추정하는 것으로 구성된다.[3]

관측 데이터 y_k 에 대하여, 최소자승법에 의하여 자기회귀(auto regression)과정을 적용한다. 즉,

$$y_k = \sum_{j=1}^p \hat{\phi}_j y_{k-j} + \rho \quad \dots (1)$$

다음, 로버스트화된 데이터 \hat{x}_k 를 다음과 같이 로버스트 필터를 통과시켜 얻는다.

$$x_k = X^T_{k-1} + c \cdot \hat{s} \cdot \phi \left[\frac{y_k - X^T_{k-1} \hat{\phi}}{c \cdot \hat{s}} \right] \quad (2)$$

$$\hat{\phi}^T = (\hat{\phi}_1 \quad \hat{\phi}_2 \quad \dots \quad \hat{\phi}_p) \quad (3)$$

$$X^T_{k-1} = (\hat{x}_{k-1} \quad \hat{x}_{k-2} \quad \dots \quad \hat{x}_{k-p}) \quad (4)$$

여기서

- \hat{s} ... innovation process r_k 의 추정 variance
- ϕ ... 로버스트화를 결정짓는 함수 (influence function)

이때 innovation process (robustly prewhitened residual), r_k 는 다음 식으로 주어진다.

$$r_k = x_k - (\hat{\phi}_1 \hat{x}_{k-1} + \hat{\phi}_2 \hat{x}_{k-2} + \dots + \hat{\phi}_p \hat{x}_{k-p}) \quad (5)$$

$$= c \cdot \hat{s} \cdot \phi \left[\frac{y_k - X^T_{k-1} \hat{\phi}}{c \cdot \hat{s}} \right]$$

이상과 같은 계산이 완료되면, x_k 는 (1)식에 대입하고, \hat{r}_k 를 이용하여 새로운 \hat{s} 를 추정한다. $\hat{\phi}$ 의 변화가 작아질 때까지 이러한 과정을 반복한다. 초기의 \hat{s} 값은 (1)식의 예측오차 e_k 의 분산의 제곱근(square root)의 값을 취한다.

최종적으로 얻어진 r_k 와 $\hat{\phi}$ 로부터 다음과 같이 로버스트화된 스펙트럼을 얻을 수 있다.

$$\bar{S}(f) = \frac{\bar{S}_r(f)}{|\bar{H}_p(f)|^2} \quad (7)$$

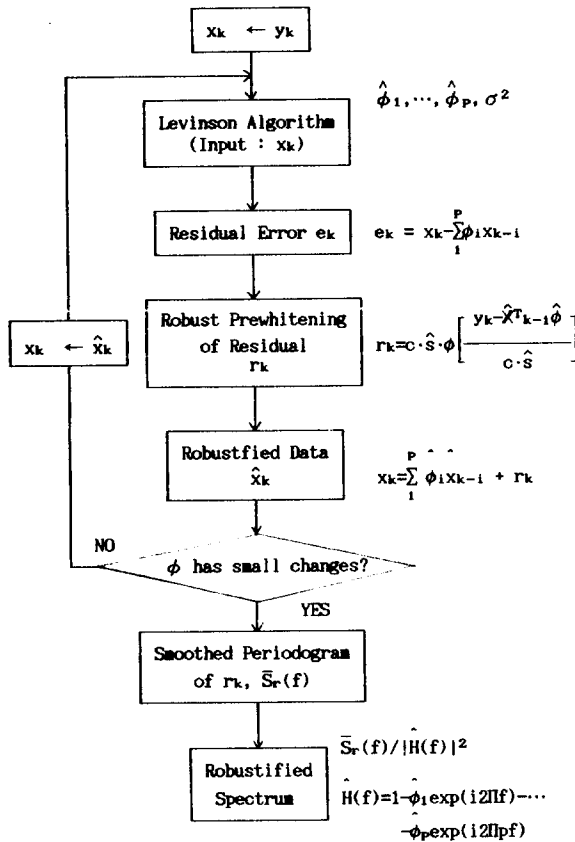
여기서, $\bar{S}_r(f)$ 는 innovation process r_k 의 averaged - periodogram을 구한 것이고 $\bar{H}_p(f)$ 는 $\hat{\phi}$ 로부터 다음 식을 이용하여 얻는다.

$$\bar{H}_p(f) = 1 - \hat{\phi}_1 \exp(i2\pi f) - \hat{\phi}_2 \exp(i2\pi 2f) - \dots - \hat{\phi}_p \exp(i2\pi pf) \quad (8)$$

이 필터는, 가우시안 분포나 꼬리가 굵은(heavy-tailed) 분포를 포함한 다양한 분포에 대하여 제곱근오차(mean-squared error)가 유한(bounded)하다는 점에서 건설 하다고 할 수 있다.

4. 실험 및 결과고찰

본 연구에서 사용한 로버스트 필터의 흐름도는 그림 4.1과 같다.



- y_k ... original data (AO model)
- x_k ... Input of AR fitting
- \hat{x}_k ... Robustified data
- e_k ... AR fitting residual (of error)
- r_k ... Robustly prewhitened residual from e_k
- \hat{S} ... variance of r_k
* at initial stage, var of e_k
- c ... efficiency coefficient (1~1.5)
- ϕ ... Influence Function

그림 4.1 로보스트 필터의 흐름도
Fig 4.1 Flow chart of robust filter

영향함수(influence function)은 건실성을 결정하는 함수인데, 주로 Huber의 psi-function 이나 Tukey의 bisquare function이 쓰인다. 이 두 함수는 각각 (9), (10)식과 같다.

Huber의 psi-function

$$\psi_n(t) = \begin{cases} t, & |t| < 1 \\ \text{sign}(t), & |t| \geq 1 \end{cases} \quad (9)$$

Tukey의 re-descending bisquare function

$$\psi_{BS}(t) = \begin{cases} t(1-t^2)^2, & |t| < 1 \\ 0, & |t| \geq 1 \end{cases} \quad (10)$$

아웃라이어에 대한 건실성을 확인하기 위하여 두 가지 모의 데이터를 발생하여 실험하였다. 첫번째 데이터 모델은 다음과 같은 세 개의 자기회귀과정(autoregressive process)으로 구성된다.[3]

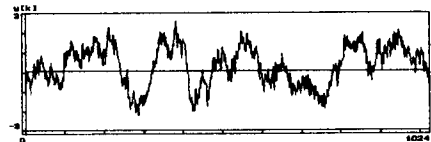
$$\begin{aligned} u_k &= 0.975 u_{k-1} + \varepsilon_k \\ w_k &= 0.95 w_{k-1} - 0.9 w_{k-2} + \eta_k \\ z_k &= 0.33 z_{k-1} - 0.9 z_{k-2} + \xi_k \end{aligned}$$

여기서 $\varepsilon_k, \eta_k, \xi_k$ 는 평균이 '0'이고 분산이 '1'인 정상 분포(normal distribution)를 갖는 백색잡음(white noise)이다. u_k 는 저주파 성분에 파워가 집중되어 있는 process이다. w_k 는 표준화 주파수 0.17에서 peak를 갖는 대역폭 0.1의 협대역 성분이며, z_k 는 표준화 주파수 0.22에서 u_k, w_k, z_k 를 각각 단위분산을 갖도록 표준화한 다음,

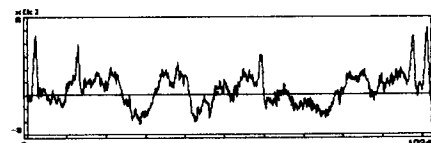
$$y_k = \sqrt{75} u_k + w_k + z_k, \quad k = 1, 2, \dots, n \quad (11)$$

을 얻고, 이를 다시 단위분산을 갖도록 표준화하여 2·AR(2) + AR(1) process를 얻는다.

그림 4.2(a)는 이렇게 하여 발생된 2·AR(2) + AR(1) process이며 그림 4.2(b)는 여기에 additive noise를 섞은 것이다. 한편, 그림 4.2(c)와 (d)는 각각 아웃라이어에 있는 경우와 없는 경우의 전력밀도 스펙트럼을, 이론적인 스펙트럼, 기존의 스펙트럼 추정 방법(본 연구에서는 AR 방법을 사용)의한 스펙트럼 그리고 로보스트 필터링 방법의 해 얻은 스펙트럼을 비교한 것이다.

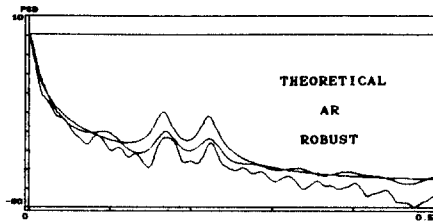


(a) 2·AR(2) + AR(1) proces

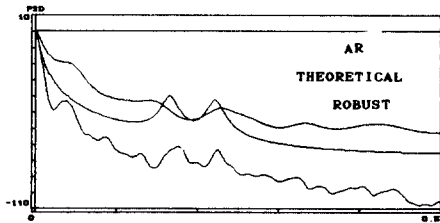


(b) 2·AR(2) + AR(1) proces + additive noise

그림 4.2 AR process



(c) 아웃라이어가 없을 때의 스펙트럼



(d) 아웃라이어가 있을 때의 스펙트럼

그림 4.2 AR process

Fig 4.2 AR process

한편 두 번째 데이터 모델로서, 다음과 같은 전달함수를 갖는 스펙트럼 형성 필터로부터 모의뇌파 신호를 발생시켰다.[4] 이 모의뇌파 스펙트럼 형성 필터는 배경뇌파의 기본 성분인 δ , α , β 리듬으로 구성된다.

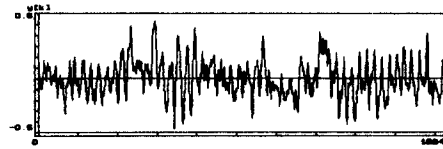
$$H(s) = -k \frac{s + 2\pi\sigma_0}{s^2 + 4\pi\sigma_p s + (2\pi f_{op})^2}$$

$$\sigma_p = \sigma_1$$

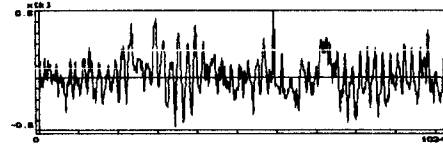
$$f_{op}^2 = (f_i^2 + \sigma_1^2)$$

$$\sigma_0^2 = (f_i^2 + \sigma_1^2) - \frac{\sigma_1 - f_i H_i / G_i}{\sigma_1 + f_i H_i / G_i}$$

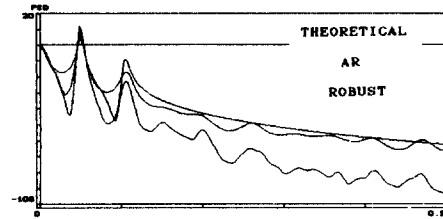
$\sigma_1 = \{2.0, 0.4, 0.8\}$, $f_i = \{0.0, 9.9, 20.9\}$, $G_i = \{55.0, 40.0, 25.0\}$ 일 때의 발생된 모의뇌파 신호는 그림 4.3(a)이고, 그림 4.3(b)는 이 모의뇌파 신호에 additive outlier를 섞은 것이다. 한편, 그림 4.3(c)와 (d)는 각각 아웃라이어가 있는 경우와 없는 경우의 전력밀도 스펙트럼을 그림 4.2와같은 방법으로 비교하여 보인 것이다.



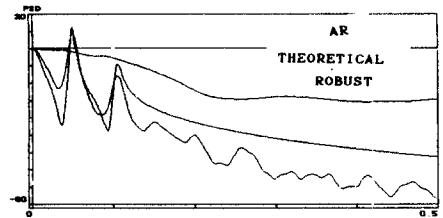
(a) 모의뇌파신호



(b) 모의뇌파신호 + additive outlier



(c) 아웃라이어가 없을 때의 스펙트럼

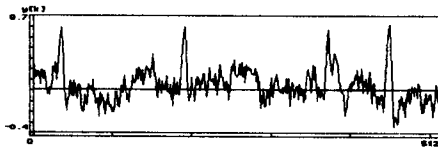


(d) 아웃라이어가 있을 때의 스펙트럼

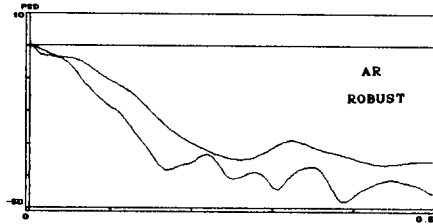
그림 4.3 모의 뇌파신호

Fig 4.3 Simulated EEG signal

그림 4.2와 그림 4.3으로부터 아웃라이어가 존재하지 않을 때는 로보스트 방법과 AR 방법에 의한 스펙트럼이 큰 차이를 보이지 않았다. 그러나 아웃라이어가 존재하는 경우에는 로보스트 방법은 스펙트럼을 비교적 정확히 추정하는 반면, AR 방법에 의해 추정된 스펙트럼에는 상당한 왜곡현상을 보이는 것을 알 수 있다. 이상의 결과로 로보스트 필터링 방법은 아티팩트에 의해 오염된 신호에 대해서 정확한 스펙트럼을 추정하는 것을 보였다. 그림 4.4는 아티팩트가 섞인 실제의 배경뇌파에 대하여 로보스트 필터링 방법을 적용해본 결과이다.



(a) 아티팩트가 섞인 배경뇌파



(b) 스펙트럼 비교

그림 4.4 실제의 배경뇌파에 대한 실험

Fig 4.4 Experiment with the real EEG

5. 결론

본 연구에서는 배경뇌파 신호의 전력밀도를 추정할 때, 혼입된 아티팩트로 인한 왜곡현상을 줄이기 위하여 로보스트 필터링 방법을 이용하였다. 이 방법은 아웃라이어(이상치)가 존재하지 않을 때는 기존의 선형적 추정방법과 같은 결과를 내면서도, 아웃라이어가 존재하더라도 스펙트럼의 왜곡현상

이 생기지 않는 건실한 방법이라 할 수 있다. 자기회귀 과정, 모의배경뇌파, 실제의 뇌파 등 여러가지 실험을 통하여 기존의 스펙트럼 추정방법과 로보스트 필터링 방법을 이용한 스펙트럼 추정방법을 비교한 결과 위와 같은 사실을 확인할 수 있었다. 따라서 로보스트 필터링 방법은 여러가지 아티팩트에 의하여 오염되기 쉬운 배경뇌파 신호는 물론이고 그 밖의 여러 생체신호의 건실한 스펙트럼을 추정하는데 적합하다고 생각된다.

6. 참고문헌

1. R. Coopee, J. W. Osselton, J. C. Shaw : EEG Technology, Butterworths, 1980
2. F.R. Hampel, E.M. Ronchetti, P.J. Rousseeuw, W.A. Stahel : Robust Statistics, 1986
3. B. Kleiner, R. D. martin, D. J. Thomson, "Robust Estimation of Power Spectra," J. R. Statist. Soc. B, 41, No3., pp.313-351
4. L. H. Zetterberg, K. Ahlin, "Analogue Simulator of EEG Signals Based on Spectral Components," Medical and Biomedical Engineering, pp.272-278, Mar. 1975
5. S. A. Kassam, H. V. Poor, "Robust Techniques for Signal processing : A Survey," Proc. of the IEEE, Vol. 73, No. 3, pp.433-481, Mar. 1985
6. D. J. Thomson, " Spectrum Estimation Techniques for Characterization and Development of WT 4 Waveguide", Bell System Techn., J. 56, pp.1769-1815, 1977